

# Testing the horizontal-vertical stereo anisotropy with the critical-band masking paradigm

**Ignacio Serrano-Pedraza**

**Claire Brash**

**Jenny C. A. Read**

Faculty of Psychology,  
Universidad Complutense de Madrid, Madrid, Spain  
Institute of Neuroscience, Newcastle University,  
Newcastle upon Tyne, UK



Imperial College London, London, UK



Institute of Neuroscience, Newcastle University,  
Newcastle upon Tyne, UK



**Stereo vision has a well-known anisotropy: At low frequencies, horizontally oriented sinusoidal depth corrugations are easier to detect than vertically oriented corrugations (both defined by horizontal disparities). Previously, Serrano-Pedraza and Read (2010) suggested that this stereo anisotropy may arise because the stereo system uses multiple spatial-frequency disparity channels for detecting horizontally oriented modulations but only one for vertically oriented modulations. Here, we tested this hypothesis using the critical-band masking paradigm. In the first experiment, we measured disparity thresholds for horizontal and vertical sinusoids near the peak of the disparity sensitivity function (0.4 cycles/°), in the presence of either broadband or notched noise. We fitted the power-masking model to our results assuming a channel centered on 0.4 cycles/°. The estimated channel bandwidths were 2.95 octaves for horizontal and 2.62 octaves for vertical corrugations. In our second experiment we measured disparity thresholds for horizontal and vertical sinusoids of 0.1 cycles/° in the presence of band-pass noise centered on 0.4 cycles/° with a bandwidth of 0.5 octaves. This mask had only a small effect on the disparity thresholds, for either horizontal or vertical corrugations. We simulated the detection thresholds using the power-masking model with the parameters obtained in the first experiment and assuming either single-channel and multiple-channel detection. The multiple-channel model predicted the thresholds much better for both horizontal and vertical corrugations. We conclude that the human stereo system must contain multiple independent disparity channels for detecting horizontally oriented and vertically oriented depth modulations.**

## Introduction

Stereo vision refers to our ability to judge depth from small disparities in the images seen by the two eyes. Because our eyes are offset horizontally in our head, these disparities are highly anisotropic, with horizontal disparities much more common than vertical ones. Depth perception, therefore, is based almost entirely on horizontal disparity. However, even when we restrict ourselves to horizontal disparities, stereo vision displays a second, puzzling anisotropy. This relates to changes in horizontal disparity (and thus depth) along horizontal or vertical directions in the image. Sinusoidal disparity corrugations at low spatial frequencies are much easier to detect when the corrugations are horizontally oriented than when they are vertically oriented (Bradshaw & Rogers, 1999; Bradshaw, Hibbard, Parton, Rose, & Langley, 2006; Serrano-Pedraza & Read, 2010; van der Willigen, Harmening, Vossen, & Wagner, 2010). Similarly, the sensitivity to disparity-defined slant is greater for surfaces rotated around the horizontal axis than for surfaces rotated around the vertical axis (Mitchison & McKee, 1990; Guillam & Ryan, 1992; Cagenello & Rogers, 1993; Hibbard, Bradshaw, Langley, & Rogers, 2002).

Recently Serrano-Pedraza and Read (2010), comparing the detectability of sinusoidal and square-wave depth corrugations, suggested that this stereo anisotropy may arise because the stereo system uses multiple spatial-frequency disparity channels for detecting horizontally oriented disparity modulations but only one for vertically oriented disparity modulations. This

Serrano-Pedraza, I., Brash, C., & Read, J. A. C. (2013). Testing the horizontal-vertical stereo anisotropy with the critical-band masking paradigm. *Journal of Vision*, 13(11):15, 1–15, <http://www.journalofvision.org/content/13/11/15>, doi:10.1167/13.11.15.

speculation was prompted by the observation that the visibility of horizontal square-wave corrugations was best predicted by the visibility of the most detectable harmonic, implying that several distinct spatial frequency channels are involved in detecting horizontal corrugations, whereas the visibility of vertical square-wave corrugations was best predicted by the root-mean-squared amplitude after filtering by a single channel. Consistent with this suggestion, the disparity sensitivity function for vertical corrugations is narrower than that for horizontal corrugations, although both peak at roughly the same value, around 0.4 cycles/°. Serrano-Pedraza and Read suggested that this could be because the disparity sensitivity function for vertical corrugations reflects only a single spatial-frequency channel, centered at 0.4 cycles/°, whereas the disparity sensitivity function for horizontal corrugations reflects contributions from multiple different channels. However, Serrano-Pedraza and Read did not carry out any experiments which directly tested for the presence of distinct channels.

Previous work has demonstrated the existence of multiple channels sensitive to horizontally oriented corrugations (Tyler, 1975; Schumer & Ganz, 1979; Tyler, 1983; Cobo-Lewis & Yeh, 1994) but until recently, no one had examined the issue for vertical corrugations. All the papers just cited used solely horizontal corrugations. Recently, Witz and Hess (2013) published the first experimental test of whether multiple channels exist for vertical disparity corrugations, using a detection  $\times$  discrimination procedure (Watson & Robson, 1981). They found that a vertical corrugation of 1 cycle/° can be discriminated from corrugations at 0.25 or 4 cycles/°, even when all three corrugations are at the threshold for detection. They concluded that there are at least three channels for vertical corrugations.

The bandwidth of disparity channels also remains unclear. Previous studies have estimated the bandwidth directly from the adaptation or masking curve. That is, the estimated bandwidth was directly taken from the threshold elevation plot, assuming that the elevation thresholds show the shape of the channel (Schumer & Ganz, 1979) or taken from the masking curves without deriving the channel tuning by means of a masking model (Cobo-Lewis & Yeh, 1994). However, these adaptation or masking curves do not necessarily reflect the bandwidth of the underlying channel (Schumer & Ganz, 1979; Cobo-Lewis & Yeh, 1994; Serrano-Pedraza, Sierra-Vázquez, & Derrington, 2013a). Unsurprisingly, therefore, the bandwidths estimated by these different techniques vary widely. Schumer and Ganz (1979), using selective adaptation, estimated the channel bandwidth at 2–3 octaves, whereas Cobo-Lewis and Yeh (1994), using a masking paradigm, estimated the channel bandwidth at 0.6 to 1.1 octaves;

both figures are for full bandwidth at half-amplitude for horizontally oriented channels. No one has estimated the bandwidth for vertical channels.

In this paper, we determine the bandwidth of the most sensitive disparity channel for both horizontal and vertical corrugations. We use the same critical-band masking paradigm as Cobo-Lewis and Yeh (1994), but rather than estimating the bandwidth directly from the masking curves, we use the classical power-spectrum model of masking developed in the study of auditory filters (Fletcher, 1940; Patterson, 1976) and here applied to stereo vision for the first time. This fitting technique enables us to achieve a more accurate estimate of channel bandwidth. We find that bandwidth is slightly narrower for vertical than horizontal corrugations. Secondly, we use a different technique to confirm Witz's and Hess's (2013) conclusion that multiple channels exist for vertical disparity corrugations. We demonstrate that disparity noise around 0.4 cycles/°, where human depth perception is most sensitive, does not impair our ability to detect corrugations two octaves lower in frequency. This implies the existence of at least two independent channels, for both horizontal and vertical depth corrugations.

## Methods

### Subjects

Experiment 1 used four human subjects: one author (ISP) and three observers unaware of the purpose of the study (JEN, RG, and PM). Experiment 2 used eight subjects: one author (ISP) and seven observers unaware of the purpose of the study (JEN, RG, PM, SP, GES, HGK, and MAMJ). All of them had experience in psychophysical experiments (aged between 18–39 years), had normal or corrected-to-normal refraction, and had normal visual acuity. Experimental procedures were approved by Newcastle University's Faculty of Medical Sciences Ethics Committee.

### Equipment and stimulus presentation

The experiments were carried out in a dark room. Stimuli were presented on a rear projection screen (300  $\times$  200 cm, Stewart Filmscreen 150, www.stewartfilm.com, supplied by Virtualis, Manchester), frontoparallel to the observers, who viewed it from a distance of 110 cm. A chin rest (UHCOTech HeadSpot) was used to stabilize the subject's head and to control the observation distance. Each eye's image was presented using a separate F20 sx+ Digital Light Processing projector

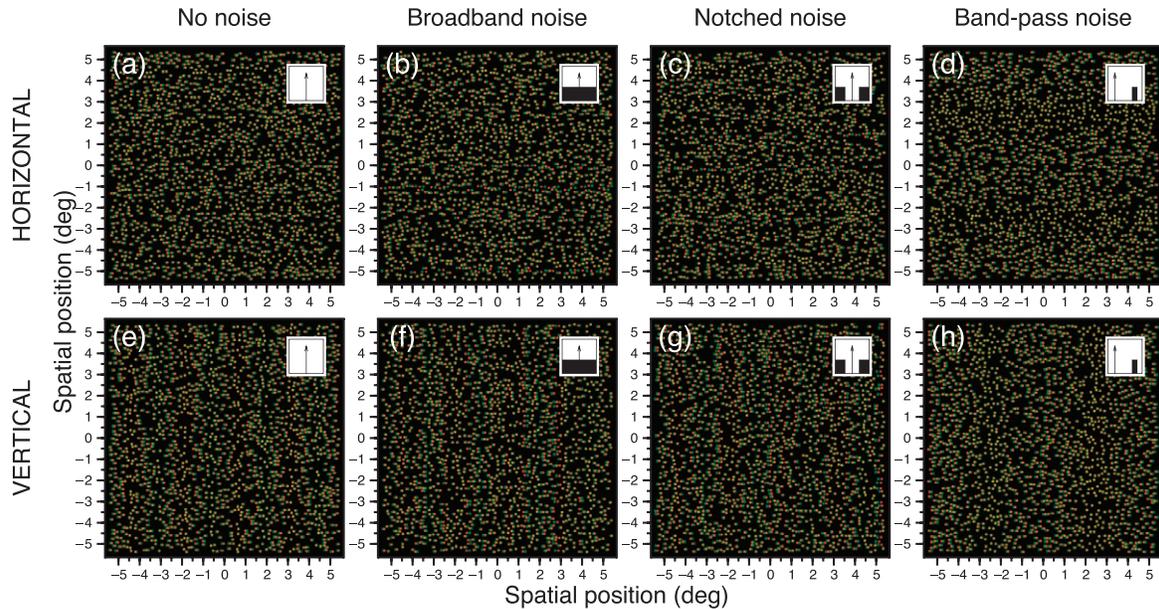


Figure 1. Anaglyph examples of random Gaussian dot stereograms used in the experiments. (a) Example of a stimulus with horizontal sinusoidal-wave corrugations of spatial frequency of  $0.4 \text{ cycles/}^\circ$  defined by horizontal disparities. (b) Same stimulus presented in (a) masked by 1D broadband noise corrugations. (c) Same stimulus presented in (a) masked by 1D notched-noise corrugations with a notch bandwidth of 3 octaves around  $0.4 \text{ cycles/}^\circ$ . (d) Example of a stimulus with horizontal sinusoidal-wave corrugations of spatial frequency of  $0.1 \text{ cycles/}^\circ$  masked by ideal 1D band-pass noise corrugations centered in  $0.4 \text{ cycles/}^\circ$  and half octave wide. In each panel there is a sketch of the amplitude spectrum of the noise used in the experiment. Panels e–f show the same stimulus but with vertical corrugations. Panels a–c and e–g show examples of stimuli used in Experiment 1. Panels d and h show examples of stimuli used in Experiment 2. [Note that the real stimuli were presented in a window of  $22^\circ \times 22^\circ$  ( $800 \times 800$  pixels) and were perceived through polarizing filters].

(ProjectionDesign, Gamle Fredrikstad, Norway; [www.projectiondesign.com](http://www.projectiondesign.com)) driven by a NVIDIA GeForce 8600 GT graphics card, with a spatial resolution of  $1400 \times 1050$  pixels (horizontal  $\times$  vertical) and a temporal resolution of 60 Hz. Both projectors were gamma corrected using a Minolta LS-100 photometer (Konica Minolta Optics, Inc., Osaka, Japan). Linear polarizing filters ensured that each eye saw only one projector's image. The cross-talk of the filters was less than 1%. The images were carefully aligned to within a pixel everywhere within the central  $30^\circ$  to ensure that as far as possible the only disparities were those introduced by the experimenter (Serrano-Pedraza & Read, 2009). The projected image had a size of  $76 \times 57$  cm subtending  $38^\circ \times 29^\circ$  (horizontal  $\times$  vertical). Each pixel thus subtended 1.6 minutes of arc (arc min). All experiments were controlled by a PC running MATLAB 7.5 (R2007b, MathWorks, Natick MA) with the Psychophysics Toolbox extensions (Brainard, 1997; Pelli, 1997; Kleiner et al., 2007; [www.psychtoolbox.org](http://www.psychtoolbox.org)). All stimuli were random-dot stereograms consisting of white dots on a black background,  $22^\circ \times 22^\circ$  ( $800 \times 800$  pixels). The dots were isotropic two-dimensional Gaussians with a standard deviation of 1.65 arc min (the dots had a dimension of  $5 \times 5$  pixels) and were scattered randomly but without overlap. The luminance of each pixel was calculated according to the value of the Gaussian

function at the center of the pixel, thus allowing subpixel effective disparities. Dot-density was  $\rho = 29.81 \text{ dots/}^\circ^2$ , giving a Nyquist limit of  $f_N = 2.73 \text{ cycles/}^\circ$   $f_N = 0.5\sqrt{\rho}$ . White on our display had a luminance of  $4 \text{ cd/m}^2$ , reduced to  $2.8 \text{ cd/m}^2$  when viewed through the polarizing glasses; the black background had a luminance of  $0.07 \text{ cd/m}^2$ , reduced to  $0.05 \text{ cd/m}^2$ . To minimize vergence movements, the subjects were instructed to maintain fixation on a small cross ( $0.3^\circ \times 0.3^\circ$ ) in the center of the screen, flanked by vertical and horizontal Nonius lines of length  $0.6^\circ$ , presented in between stimuli. In each experiment, a new trial was initiated after the participant's response, thus the experiments proceeded at a pace determined by the observer. No feedback about correctness on individual trials was given.

## Stimulus construction

### Signal depth corrugations

In both experiments, the task was to detect signal corrugations. These were sinusoidal depth corrugations defined by horizontal disparity, oriented either horizontally or vertically (see Figures 1a and 1e). In Experiment 1, the spatial frequency of the signal corrugation was  $0.4 \text{ cycles/}^\circ$ . This was presented either unmasked, or masked by either broadband or notched

one-dimensional (1D) Gaussian noise (see Figures 1b, 1c, 1f, and 1g). In Experiment 2, the spatial frequency of the signal corrugation was 0.1 cycles/°, and it was masked by band-pass 1D noise centered on 0.4 cycles/°. In each case, the noise was 1D with the same orientation as the signal. Thus for example for horizontal gratings, the disparity of the noise was constant along each row of pixels.

### Amplitude spectra of the noise

Our experimental apparatus enabled us to display disparities with spatial frequencies between 0.04 cycles/° and 2.7 cycles/°. The broadband noise used in Experiment 1 had a flat amplitude spectrum between the limits 0.04–2.5 cycles/°. We used broadband noise with five different power spectral density or noise levels  $N_0 \in \{1.625, 6.5, 25, 100, 400\} \times 10^{-3}$  (cycles/°)<sup>-1</sup>.

For the notched noise, the amplitude spectrum was flat except for a “notch,” centered on 0.4 cycles/°, where the stimulus had no power. We used notch bandwidths of  $S_{oct} = 0.5, 1, 2, 3,$  and 4 octaves; notches were symmetrical in log-frequency. Thus for a 1-octave notch, the noise amplitude spectrum was zero between 0.28 and 0.57 cycles/°. For the notch stimuli, the noise level outside the notch was always  $N_0 = 25 \times 10^{-3}$  (cycles/°)<sup>-1</sup>, independent of the bandwidth of the notch.

In Experiment 2, we used band-pass noise centered on 0.4 cycles/° with a spatial-frequency bandwidth of 0.5 octaves (see Figures 1d and 1h). The amplitude spectrum was flat between 0.34 to 0.48 cycles/°, and zero outside this range. The noise level was again  $N_0 = 25 \times 10^{-3}$  (cycles/°)<sup>-1</sup>.

### Generating a noise sample

The horizontal disparities of the 1D noise masks were calculated in the Fourier domain following the same steps that are usually used to construct luminance noises (see a detailed procedure for luminance 1D white noise in Serrano-Pedraza & Sierra-Vázquez, 2006). First, we generated the desired amplitude spectrum, as described above. Second, we generated a phase spectrum from random variables uniformly distributed between  $(-\pi, \pi]$ . Third, we transformed the Fourier spectrum into the spatial domain. In this way, we generated a sample of 1D disparity noise which can then be displayed with the desired amplitude or noise level. We then converted this to a 2-D disparity map  $n(x, y)$  by extending the 1D disparity noise in the orthogonal direction.

### Setting the noise level

In order to present a noise sample with a particular noise level we first calculated the desired root mean square amplitude of the disparity noise using the

equation  $V_{RMS} = \sqrt{N_0 2W}$ , (Green & Swets, 1974) where  $N_0$  is the noise level and  $W$  is the bandwidth of the noise that had energy (e.g., for our broadband noise  $W$  was 2.46 cycles/°). Second, we normalized the noise sample to peak at one. Third, in order to present the sample of disparity noise with the desired  $V_{RMS}$ , we calculated the peak amplitude disparity ( $A$ ) using the following equation:

$$A = \frac{V_{RMS}}{\sqrt{a - b^2 - bV_{RMS}}}, \quad (1)$$

where  $a = \{[\sum \sum n^2(x, y)]/MN\}$ ,  $b = \{[\sum \sum n(x, y)]/MN\}$ , and  $MN$  is the number of columns ( $M$ )  $\times$  rows ( $N$ ) of the noise sample (800  $\times$  800 pixels in the experiments). Finally, we multiplied the sample of disparity noise ( $n(x, y)$ ) by the desired disparity amplitude ( $A$ ), and added the result to the signal in order to obtain the desired disparity map (see Figures 1b and 1f). Note that  $V_{RMS}$  is the standard deviation of the result of multiplying  $A \times n(x, y)$  (Davenport & Root, 1958).

### Generating the random-dot pattern

Finally, we displayed the resulting disparity map via a random-dot pattern. We generated around 14,000 dots scattered randomly across the two-dimensional image. The position of each dot was rounded to the nearest pixel; then we looked up from the disparity map what the desired disparity was at this point and shifted the horizontal positions of the dot in left and right eyes accordingly.

### Procedure

Peak amplitude disparity thresholds for unmasked and masked sinusoidal disparity corrugations were measured using adaptive Bayesian staircases (Treutwein, 1995) in a two-interval forced-choice task. For unmasked conditions, random-dot stereograms were presented with zero disparity in one presentation interval and with a disparity corrugation in the other. The task was to indicate the interval containing the disparity corrugation. For masked conditions, the 1D disparity noise (mask) was presented in both presentation intervals, and the sinusoidal corrugations (signal) was added to the mask in one presentation interval. The task was to indicate the interval containing the sinusoidal corrugation (signal). A different uniform random distribution of dots was presented in each interval and a different sample of noise was presented in each trial. Corrugation orientation (vertical or horizontal) were blocked, as was

noise type (broadband or notch) in Experiment 1, but the noise parameters (noise level for broadband noise and notch bandwidth for notched noise) were interleaved.

Each presentation interval was preceded by a vertical Nonius line presented for 300 ms followed by 200 ms of a blank screen. The presentation intervals lasted 250 ms, so the total trial duration was 1500 ms (Nonius line + blank + first interval + Nonius line + blank + second interval). In general between 6 and 8 min were required per disparity threshold estimation. The characteristics of the Bayesian staircases were: (a) the prior probability density function was uniform (Pentland, 1980; Emerson, 1986); (b) the model likelihood function was the logistic function adapted from García-Pérez (1998, his appendix A) with a spread value of 0.8 (with delta parameter equal to 0.01), a lapse rate of 0.01, and a guess rate of 0.5; (c) the value of the disparity in each trial was obtained from the mean of the posterior probability distribution (King-Smith, Grigsby, Vingrys, Benes, & Supowit, 1994); (d) the staircase stopped after 50 trials (Pentland, 1980; Anderson, 2003); and (e) the final threshold was estimated from the mean of the final probability density function. The disparity threshold corresponded to the value 0.85 of the subject's psychometric function. Three threshold estimations per condition were obtained for each subject. In Experiment 1, a total of 42 conditions (1 Unmasked Sinusoidal Corrugation of Frequency 0.4 cycles/° × 2 Orientations; 1 Masked Sinusoidal Corrugation of Frequency 0.4 cycles/° × 5 Noise Levels of Broadband Noise × 4 Notch Bandwidths × 2 Orientations) were tested in different experimental sessions. In Experiment 2, a total of four conditions were tested (1 Unmasked Sinusoidal Corrugation of Frequency 0.1 cycles/° × 2 Orientations; 1 Masked Sinusoidal Corrugation of Frequency 0.1 cycles/° × 1 Band-pass Noise Centered in 0.4 cycles/° × 2 Orientations).

## The power-spectrum model of visual masking

In order to explain the stereo masking results we will use the power-spectrum model of masking. This classical model, adapted from the study of auditory filters (Fletcher, 1940; Patterson, 1976; Patterson & Nimmo-Smith, 1980; Patterson & Moore, 1986, p. 124; Moore, 1997), is one of the most used in vision to study luminance spatial-frequency channels (Pelli, 1981; Perkins & Landy, 1991; Solomon & Pelli, 1994; Losada & Mullen, 1995; Blackwell, 1998; Mullen & Losada, 1999; Solomon, 2000; Majaj, Pelli, Kurshan, & Palomares, 2002; Taltar, Pelli, & Carrasco, 2004; Serrano-Pedraza & Sierra-Vázquez, 2006; Serrano-Pedraza, Sierra-Vázquez, et al., 2013a; Westrick,

Henry, & Landy, 2013). In this work we will adapt this model to explain stereo masking results and to study the characteristics of the visual disparity channels.

The model makes three main assumptions:

- Stimuli are processed by a bank of separate, independent, and overlapping band-pass linear channels, each tuned to a different spatial frequency  $\xi$ ;
- A channel  $k$  detects a signal when its power signal-to-noise ratio reaches some fixed threshold  $\theta$ ; and
- A channel's sensitivity is limited by its internal noise  $N(\xi_k)$ .

The threshold  $\theta$  sets the overall sensitivity of the system,  $s = 1/\theta$ , while  $N(\xi)$  sets the relative sensitivity between channels. The channel's power response to a grating of disparity amplitude  $m$  at spatial frequency  $u_0$  is  $\{[m^2(u_0)]/2\}|H(u_0; \xi_k)|^2$ , where  $H(u, \xi)$  is the channel's modulation transfer function (see Equation 4), normalized such that  $H(\xi_k, \xi_k) = 1$ . The channel's power response to 1D noise with power spectrum  $\rho(u)$  is  $2 \int_0^{+\infty} \rho(u) |H(u; \xi_k)|^2 du$ . Taking the internal noise into account, and following Assumption 2, the power signal-to-noise ratio for the minimum amplitude  $m_T(u_0)$  necessary for the channel  $k$  to detect a signal at  $u_0$  in the presence of 1D noise is

$$\frac{\frac{m_T^2(u_0)}{2} |H(u_0; \xi_k)|^2}{N(\xi_k) + 2 \int_0^{+\infty} \rho(u) |H(u; \xi_k)|^2 du} = \frac{1}{s}. \quad (2)$$

In the absence of external noise, the disparity threshold  $m_0$  at any frequency is set by the internal noise of the channel detecting it. A channel's internal noise can therefore be deduced from the sensitivity at the channel's preferred frequency:

$$N(\xi_k) = sm_0^2(\xi_k)/2.$$

Substituting this expression for internal noise into Equation 2, we obtain the fundamental masking equation:

$$m_T^2(u_0) = \frac{m_0^2(\xi_k) + \frac{4}{s} \int_0^{+\infty} \rho(u) |H(u; \xi_k)|^2 du}{|H(u_0; \xi_k)|^2}, \quad (3)$$

which relates the increased threshold needed to detect a masked signal to the unmasked threshold and, critically, the channel modulation transfer function (MTF) (Serrano-Pedraza & Sierra-Vázquez, 2006; Serrano-Pedraza, Sierra-Vázquez, et al., 2013a). Equation 3 was first derived, with minor differences, for luminance channels by Solomon (2000, see his equations 4 and 5).

### MTF of the disparity channels

As MTF we used the log-Gaussian function (Morrone & Burr, 1988):

$$|H(u; \xi_i)| = \begin{cases} \exp\left[-\frac{\ln^2(|u|/\xi_i)}{2\alpha_i^2}\right] & \Leftrightarrow u \neq 0 \\ 0 & \Leftrightarrow u = 0 \end{cases}, \quad (4)$$

where  $\xi_i$ ,  $\xi_i \neq 0$ , is the peak spatial frequency of the disparity channel  $i$  and  $\alpha_i$ ,  $\alpha_i > 0$ , is an index of its spatial spread. The relative bandwidth (full width at half height), in octaves, is obtained from  $B_{oct}(\xi_i) = (2\sqrt{2}/\sqrt{\ln 2})\alpha_i$ . We have chosen this MTF because: (a) we know the analytic solution of its integral, which will be useful for fitting the power-spectrum model (see Appendix, Equation A1); and (b) it has a symmetric shape when represented in log scale, similar to the adapting thresholds curves found with disparity gratings (Schumer & Ganz, 1979).

### Detection models

The power-spectrum model assumes that the visual channel that detects the signal is the one with the highest ratio of signal power to noise power at its output. Under some circumstances, e.g., if the noise is broadband, this will be the channel most closely tuned to the signal. Under other circumstances, a different channel may have the highest signal-to-noise ratio (off-frequency looking) (Patterson & Nimmo-Smith, 1980; Pelli, 1981; Solomon, 2000; Serrano-Pedraza, Sierra-Vázquez, et al., 2013a).

In Experiment 1 we use broadband noise (white noise) and notched noise as masks. Previous studies have shown that these noise profiles prevent off-frequency looking (Serrano-Pedraza & Sierra-Vázquez, 2006; Serrano-Pedraza, Sierra-Vázquez, et al., 2013a). Because our masks prevents off-frequency looking, we can assume that sinusoidal depth corrugations of spatial frequency  $u_0$  are always detected by the disparity channel which is tuned most closely to the spatial frequency of the signal. We know that there must be a channel tuned to 0.4 cycles/°, since that is the peak of the disparity sensitivity function (Rogers & Graham, 1982; Bradshaw & Rogers, 1999; Serrano-Pedraza & Read, 2010). We can therefore assume that  $\xi_k = u_0$ ,  $u_0 = 0.4$  cycles/° (see Appendix and Equations A2 and A3).

Thus, our modeling for Experiment 1 does not depend on whether disparity corrugations are detected by multiple channels or by only one. If there is only one channel, as postulated by Serrano-Pedraza and Read (2010) for vertical corrugations, it must be at 0.4 cycles/°. If there are multiple channels, as we know there are for horizontal corrugations, they do not affect the results since the model predicts that only the channel

centered on the spatial frequency of the signal, 0.4 cycles/°, is the one that will detect the signal. In Experiment 2, we model different predictions based on which channel detects the signal.

In Experiment 2, we compare the different predictions for single versus multiple channels. Here, we used band-pass noise centered on 0.4 cycles/° and 0.5 octaves wide. The spatial frequency of the sinusoidal depth corrugation was  $u_0 = 0.1$  cycles/°. For the single-channel hypothesis, of course, there is only a single channel, tuned to  $\xi_k = 0.4$  cycles/° (see Appendix and Equation A4). For the multiple-channel hypothesis, we need to consider which channel would detect the stimulus. With band-pass noise, off-frequency looking becomes possible: the stimulus can be detected by a disparity channel tuned to a spatial frequency different from that of the signal. Serrano-Pedraza et al. (2013a) showed that when the center spatial frequency of the band-pass noise is situated more than 2 octaves of distance from the spatial frequency of the signal then there is almost no off-frequency looking. Here, assuming that all disparity channels of a given orientation have the same bandwidth (Schumer & Ganz, 1979) and taking this to be the value obtained in Experiment 1, we have found by simulation, following the procedure described in Serrano-Pedraza et al. (2013a), that the channel with the highest signal to noise ratio for detection of the signal is tuned to 0.09 cycles/°. Since this is so close to 0.1 cycles/°, for the multiple disparity channels hypothesis we have assumed that the channel that detects the signal is tuned to  $\xi_k = 0.1$  cycles/°.

### Fitting the power-spectrum model to the data and model predictions

We fitted the power-spectrum model to the data of Experiment 1. We used Equation A2 for broadband masking data and Equation A3 for notched noise masking data. The model has two parameters:  $\alpha_i$  and  $s$ . Parameter  $\alpha_i$  controls the bandwidth ( $B_{oct}$ , in octaves) of the disparity channel and  $s$  its sensitivity. For each subject we fitted Equations A2 and A3 to the data, where the values of the parameters  $\alpha_i$  and  $s$  were estimated using a least-squares fitting procedure. The sum of the squared errors between the empirical squared-disparity thresholds and the model squared-disparity thresholds was minimized using the Matlab routine “fminsearch” that uses the Nelder-Mead simplex search method (Nelder & Mead, 1965). The goodness of fit was calculated by means of the coefficient of determination ( $R^2$ ) between the model predictions and all empirical masking thresholds (broadband and notched noise masking data) (see red line in Figures 2, 3, and 4).

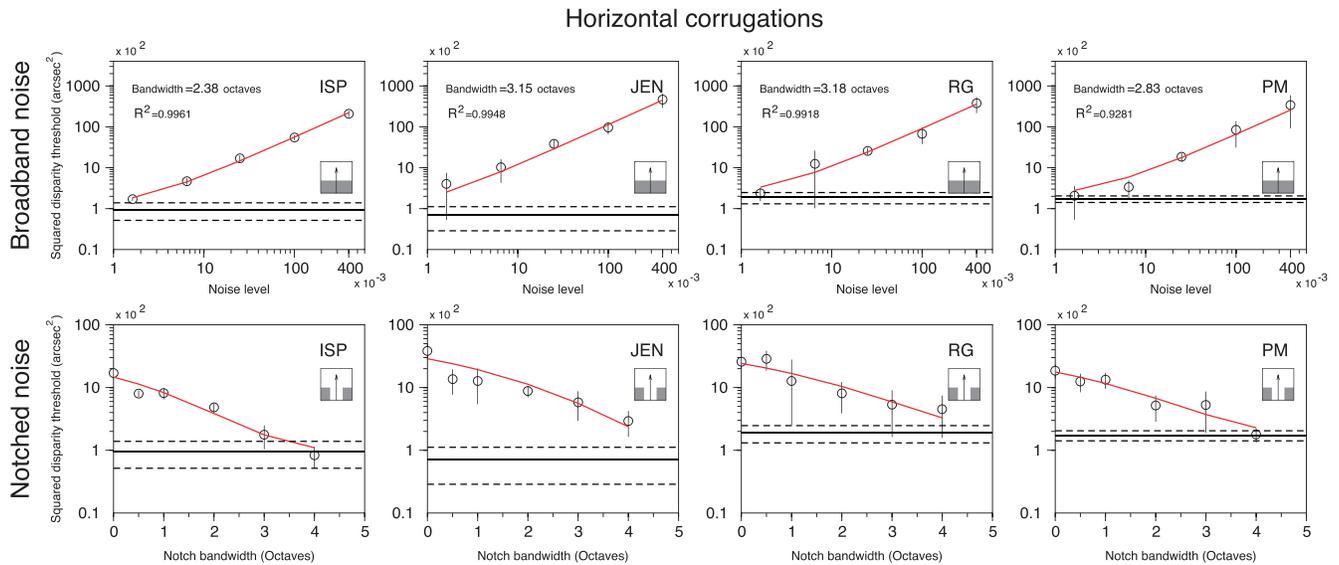


Figure 2. Masking results from four subjects for horizontally oriented depth corrugations. Each column shows the results of one subject. Upper panels show the squared disparity thresholds (mean  $\pm$  SD) for a sinusoidal corrugation of spatial frequency of 0.4 cycles/° as a function of the masking noise level (in units (cycles/°)<sup>-1</sup>) of the broadband noise. Lower panels show the squared disparity thresholds (mean  $\pm$  SD) for a sinusoidal corrugation of 0.4 c/° as a function of the notched noise bandwidth (in octaves). In each panel there is a sketch of the amplitude spectrum of the noise used in the experiment. The noise level (power spectral density) for the notched noise was  $25 \times 10^{-3}$  (cycles/°)<sup>-1</sup>. The red line shows the fitting of the power spectrum model. The top panels show the estimated value of the bandwidth in octaves. The value of  $R^2$  shown in the top panels of each row is the coefficient of determination between all masking thresholds from the two conditions (broadband and notched noise) and the model predictions. The shape of the channel was the lognormal function (see text for details).

The objective of Experiment 2 was to compare the hypotheses that variations in disparity are detected by single or multiple disparity channels. We used the estimated values  $\alpha_i$  and  $s$  obtained from Experiment 1 for vertical and horizontal corrugations (see Figure 4). The predictions were the detection thresholds for detecting a sinusoidal corrugation of 0.1 cycles/° (vertical and horizontal corrugations), either unmasked or masked by band-pass noise centered at 0.4 cycles/° (0.5 octaves width) (see example in Figures 1d and 1h). To obtain predictions under the two hypotheses, we run the model (see Appendix and Equation A4) assuming that the signal is detected by a disparity channel centered at spatial frequency of either 0.1 cycles/° (multiple channel hypothesis) or 0.4 cycles/° (single channel hypothesis).

## Results

### Experiment 1. Masking horizontal and vertical corrugations using broadband noise and notched noise

The objective of this experiment was to determine the bandwidth of the disparity channel tuned to 0.4

cycles/°. We used this spatial frequency because it has been reported previously that the minimum disparity threshold (maximum sensitivity) is obtained for a corrugation frequency about 0.4 cycles/° (Rogers & Graham, 1982; Tyler, 1991; Bradshaw & Rogers, 1999) for both horizontal and vertical corrugations (Serrano-Pedraza & Read, 2010). This indicates that the stereo system has, as a minimum, a channel centered on or near 0.4 cycles/°; of course, other channels may exist as well. The result of this experiment will be used in Experiment 2 in order to predict the effect of noise at 0.4 cycles/° under different assumptions. In Experiment 1 we measured disparity thresholds for sinusoidal depth corrugations masked by depth corrugations of broadband noise and notched noise. We examined both horizontal and vertical orientations (see example stimuli in Figures 1b, 1c, 1f, and 1g).

Figure 2 shows the masking results for four subjects for horizontal corrugations; Figure 3 shows the same for vertical. Top panels show the results for broadband masking noise and bottom for notched masking noise. Each panel shows the squared disparity thresholds (open circles) for detecting a sinusoidal depth corrugation of spatial frequency of 0.4 cycles/° as a function of the noise level (top panels) or as a function of the notch bandwidth (bottom panels). The black horizontal line shows the squared disparity threshold for the sinusoidal depth corrugation without masking. As

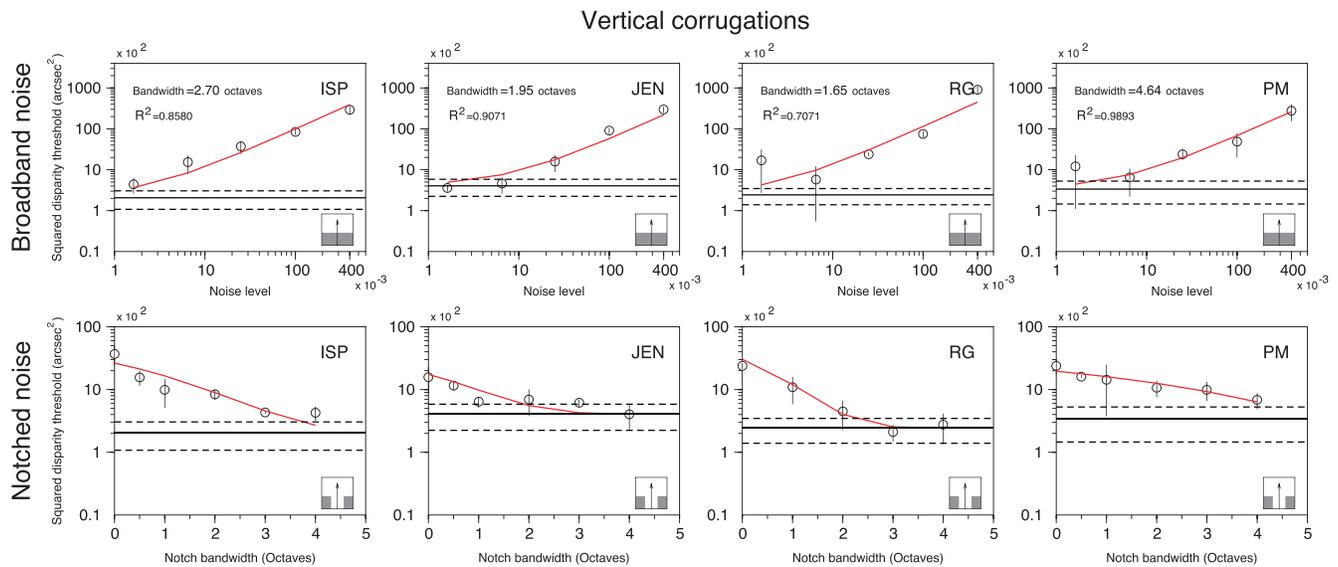


Figure 3. Masking results from four subjects for vertically oriented depth corrugations. Each column shows the results of one subject. Upper panels show the squared disparity thresholds (mean  $\pm$  SD) for a sinusoidal corrugation of spatial frequency of  $0.4 \text{ cycles/}^\circ$  as a function of the masking noise level (in units  $(\text{cycles/}^\circ)^{-1}$ ) of the broadband noise. Lower panels show the squared disparity thresholds (mean  $\pm$  SD) for a sinusoidal corrugation of  $0.4 \text{ cycles/}^\circ$  as a function of the notched noise bandwidth (in octaves). In each panel there is a sketch of the amplitude spectrum of the noise used in the experiment. The noise level (power spectral density) for the notched noise was  $25 \times 10^{-3} (\text{cycles/}^\circ)^{-1}$ . The red line shows the fitting of the power spectrum model. The value of  $R^2$  shown in the top panels of each row is the coefficient of determination between all masking thresholds from the two conditions (broadband and notched noise) and the model predictions. The top panels show the estimated value of the bandwidth in octaves. The shape of the channel was the lognormal function (see text for details).

expected, when broadband disparity noise is used as the mask, the disparity thresholds increase with the increasing noise level (see Figures 2 and 3, top panels), whereas, when notched noise is used as the mask, the disparity thresholds decrease with the increasing bandwidth of the notch (see Figures 2 and 3, bottom panels).

The data in Figures 2 and 3 are similar to the data obtained in luminance studies (Pelli, 1981; Losada & Mullen, 1995; Mullen & Losada, 1999; Solomon, 2000; Serrano-Pedraza & Sierra-Vázquez, 2006; Serrano-Pedraza, Sierra-Vázquez, et al., 2013a). The masking data for horizontally oriented depth corrugations replicate those obtained by Cobo-Lewis and Yeh (1994) (see their figures 4, 6, and 7).

The red curves of Figures 2 and 3 show the predictions of the power-spectrum model of visual masking with the fitted parameters. The predictions for broadband noise (top panels) are given by Equation A2 and for notched noise (bottom panels) by Equation A3. We fitted both experimental conditions (broadband and notched noise) together, so fitting two free parameters to a total of 11 data points. The full-width half-amplitude bandwidth of the disparity channel estimated from the fitting is specified in the top panels for each subject. The coefficient of determination ( $R^2$ ) between all masking thresholds and the model predic-

tion is specified in the top panels of Figures 2 and 3. For horizontal corrugations (Figure 2), the estimated bandwidths ranged from 2.4 to 3.2 octaves in our four subjects, with a mean of 2.9 octaves. For vertical corrugations, estimated bandwidths ranged from 1.7 to 4.6 octaves, with a mean of 2.7 octaves (see Figure 3).

Figure 4 shows the average of the disparity thresholds of the four subjects from Figures 2 and 3. We fitted the power-spectrum model to this mean data as we did in Figures 2 and 3 for the individual subjects. As described in Figures 2 and 3, the black horizontal line shows the squared disparity threshold (mean  $\pm$  SD of four subjects) for the sinusoidal depth corrugation without masking. The average of the disparity thresholds for horizontal corrugations was 11.5 arcsec, and for vertical corrugations was 17.24 arcsec (vertical/horizontal [V/H] ratio is 1.5). This is similar to previous estimates of the relatively weak stereo anisotropy at this frequency (Bradshaw & Rogers, 1999; Serrano-Pedraza & Read, 2010).

We estimated the disparity channel bandwidth (full-width half-amplitude bandwidth) for both orientations: For horizontal corrugations we found a bandwidth of 3.0 octaves and for vertical corrugations we found a bandwidth of 2.6 octaves. These values are close to the means of the values for the individual subjects.

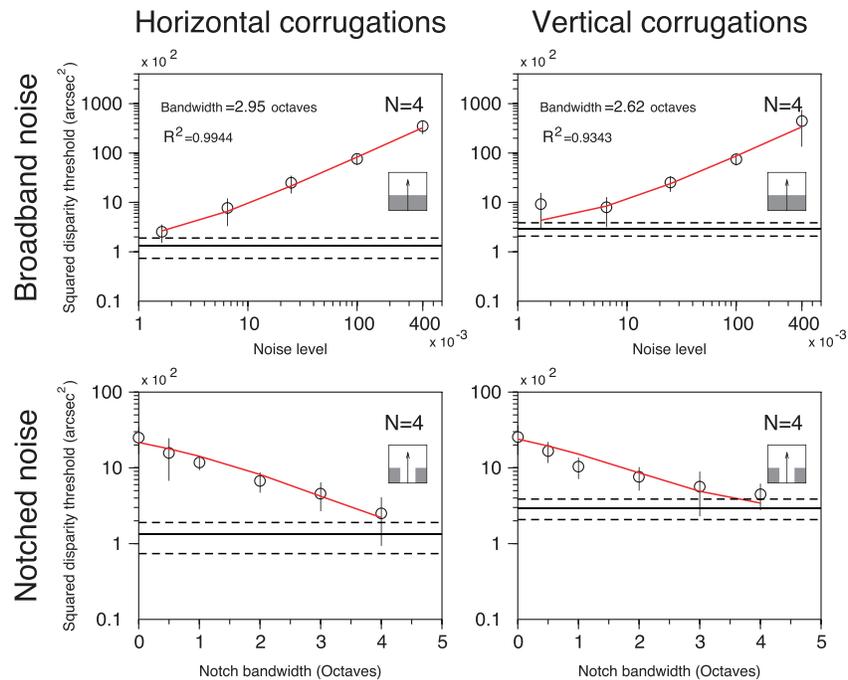


Figure 4. Average masking results from Figures 2 and 3 for horizontally (left column) and vertically oriented (right column) depth corrugations. Upper panels show the mean ( $\pm$  SD) of the squared disparity thresholds for a sinusoidal corrugation of spatial frequency of  $0.4 \text{ cycles}/^\circ$  as a function of the masking noise level (in units  $(\text{cycles}/^\circ)^{-1}$ ) of the broadband noise. Lower panels show the squared disparity thresholds (mean  $\pm$  SD) for a sinusoidal corrugation of  $0.4 \text{ cycles}/^\circ$  as a function of the notched noise bandwidth (in octaves). In each panel there is a sketch of the amplitude spectrum of the noise used in the experiment. The noise level (power spectral density) for the notched noise was  $25 \times 10^{-3} (\text{cycles}/^\circ)^{-1}$ . The red line shows the fitting of the power spectrum model. The value of  $R^2$  shown in the top panels of each row is the coefficient of determination between all masking thresholds from the two conditions (broadband and notched noise) and the model predictions. The top panels show the estimated value of the bandwidth in octaves. The shape of the channel was the lognormal function (see text for details).

## Experiment 2. Masking horizontal and vertical corrugations using band-pass noise

The objective of this experiment was to test if vertically oriented depth corrugations are detected by a single or by multiple disparity channels. In Experiment 2 we measured disparity thresholds for sinusoidal depth corrugations defined by horizontal disparity of spatial frequency of  $0.1 \text{ cycles}/^\circ$ . We compared thresholds without noise with those obtained with band-pass noise centered on  $0.4 \text{ cycles}/^\circ$  (see Figures 1d and 1h). As before, we tested horizontal and vertical corrugations.

Figure 5 shows the results for eight subjects. The left column shows the results for horizontal corrugations, while the right column shows the results for vertical corrugations. The top panels show the disparity thresholds (in seconds of arc). Green dots show the disparity thresholds for detecting the sinusoidal depth corrugation without masking. Thresholds are generally higher for the vertical corrugations, reflecting the well-known stereo anisotropy found at low spatial frequencies (Bradshaw & Rogers, 1999; Serrano-Pedraza & Read, 2010). The average of the disparity thresholds

for horizontal corrugations was  $22.6 \text{ arcsec}$  and for vertical corrugations was  $51.54 \text{ arcsec}$  (V/H ratio is 2.27). Red squares show the disparity thresholds for detecting the sinusoidal depth corrugation masked by band-pass noise. The black dots in the bottom panels of Figure 5 show the ratio of the masked to nonmasked thresholds for each subject. The ratios are close to one, meaning that noise at  $0.4 \text{ cycles}/^\circ$  has little effect on subjects' ability to detect corrugations at  $0.1 \text{ cycles}/^\circ$ , even though we are much more sensitive to disparity at  $0.4 \text{ cycles}/^\circ$ . This already enables us to conclude that the signals at  $0.1$  and  $0.4 \text{ cycles}/^\circ$  are detected by different channels and thus that at least two channels exist for both horizontal and vertical corrugations.

To quantify this, we used the power-spectrum model with the bandwidths obtained from the average data of the Experiment 1 (Figure 3). Our model assumes that the internal noise of each channel is given by the disparity threshold at that channel's peak spatial frequency. Three such thresholds are shown in Figure 6a. Figure 6b shows the predicted masked thresholds as a function of the channel assumed to be detecting signals at  $0.1 \text{ cycles}/^\circ$ , using the noise implied by Figure

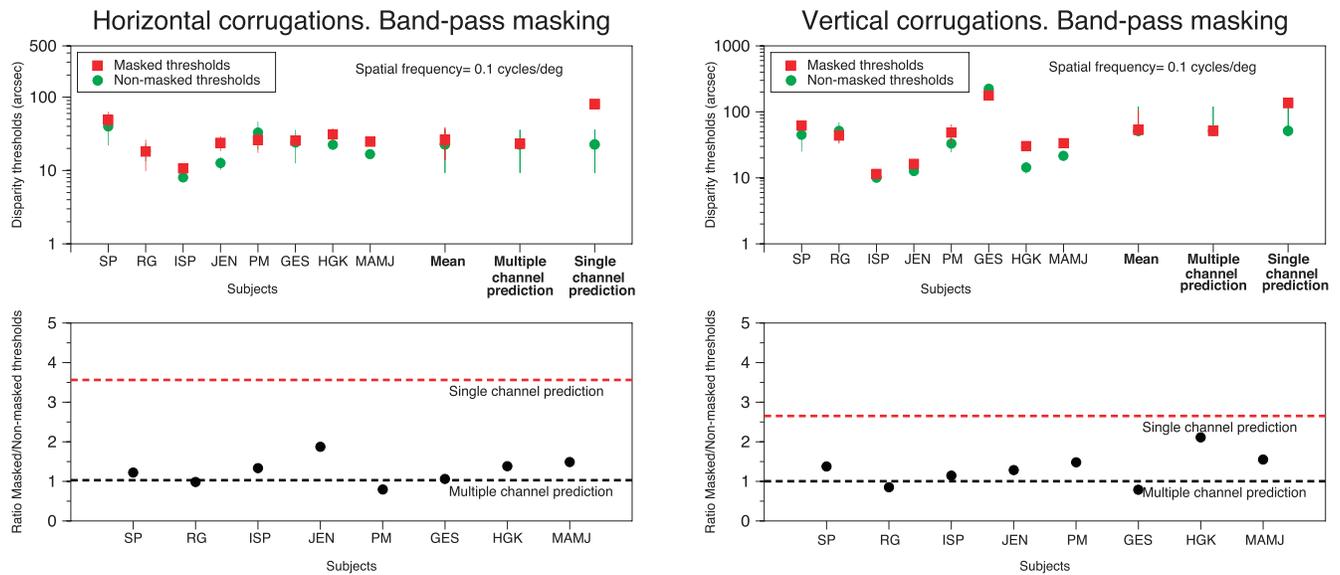


Figure 5. Masking results from Experiment 2 for horizontally (left column) and vertically oriented (right column) depth corrugations. Upper panels show: (a) green dots, the disparity thresholds for a sinusoidal corrugation of spatial frequency of  $0.1 \text{ cycles/}^\circ$ ; (b) red squares, disparity thresholds of a sinusoidal corrugation of spatial frequency of  $0.1 \text{ cycles/}^\circ$  masked by ideal 1D band-pass noise centered in  $0.4 \text{ cycles/}^\circ$  and half octave wide. The noise level (power spectral density) for the band-pass noise was  $25 \times 10^{-3} (\text{cycles/}^\circ)^{-1}$ . The upper panels also show the mean result of eight subjects (mean  $\pm$  SD) and the predictions (red squares) of the power-masking model (assuming visual channels with spatial-frequency bandwidths from Figure 4, see text for details) for single and multiple disparity channels. Lower panels (black dots) show the ratio of the masked thresholds (red squares) versus nonmasked thresholds (green dots) for each subject. The black dashed line shows the ratio for the multiple channel prediction. The red dashed line shows the ratio for the single channel prediction.

6a (interpolating for  $0.3 \text{ cycles/}^\circ$ ) and assuming that all channels have the same bandwidth (that measured in Experiment 1). Figure 6c shows the same results expressed as a ratio of masked to unmasked thresholds. If the channel at  $0.4 \text{ cycles/}^\circ$  were the only channel present, then this would have to be used for detecting signals at  $0.1 \text{ cycles/}^\circ$ , in which case noise at  $0.4 \text{ cycles/}^\circ$  would be predicted to elevate vertical detection thresholds by a factor of 2.5 (black dot at  $0.4 \text{ cycles/}^\circ$  in Figure 6c, or red dashed line in Figure 5, right panel). If there are channels present at all frequencies, we can assume the signal is detected by a channel close to  $0.1 \text{ cycles/}^\circ$ , in which case noise at  $0.4 \text{ cycles/}^\circ$  would have no effect on detected thresholds (dots at  $0.1 \text{ cycles/}^\circ$  in Figure 6c). Clearly, this is much closer to what we observe (dashed lines in Figures 6b and c). Our modeling enables us to conclude that there are at least two channels present for each orientation, one at the peak sensitivity,  $0.4 \text{ cycles/}^\circ$ , and one at a lower frequency, no greater than  $0.2 \text{ cycles/}^\circ$ .

## Discussion

In Experiment 1, we determined the bandwidth of the most sensitive disparity channel, the one centered

on  $0.4 \text{ cycles/}^\circ$ . We measured disparity thresholds for detecting a sinusoidal disparity corrugation of  $0.4 \text{ cycles/}^\circ$  oriented vertically and horizontally and under unmasked and masked conditions. In the masked condition we used 1D broadband noise with different noise levels and 1D notched noise with different spectral gaps around the spatial frequency of the signal. By fitting the power-spectrum model of visual masking to the masking data (see red line in Figures 2, 3, and 4), we estimated full-width half-amplitude bandwidth at around 3.0 octaves for horizontal corrugations and similar but slightly smaller for vertical corrugations at around 2.6 octaves.

Cobo-Lewis and Yeh (1994) measured disparity thresholds for horizontally oriented sinusoidal depth corrugations masked by notched and narrowband disparity noises. They reported masking curves of 1.1 octaves bandwidth (full-width half-amplitude bandwidth) when using notched noise, and masking curves of 0.6 octaves (full-width half-amplitude bandwidth) when using narrowband noise. The authors did not use a masking model to interpret their data, so they estimated the bandwidths directly from the masking curves, not from the underlying disparity channels that detect the signals. As these authors concede, this approach can give narrower bandwidths than the

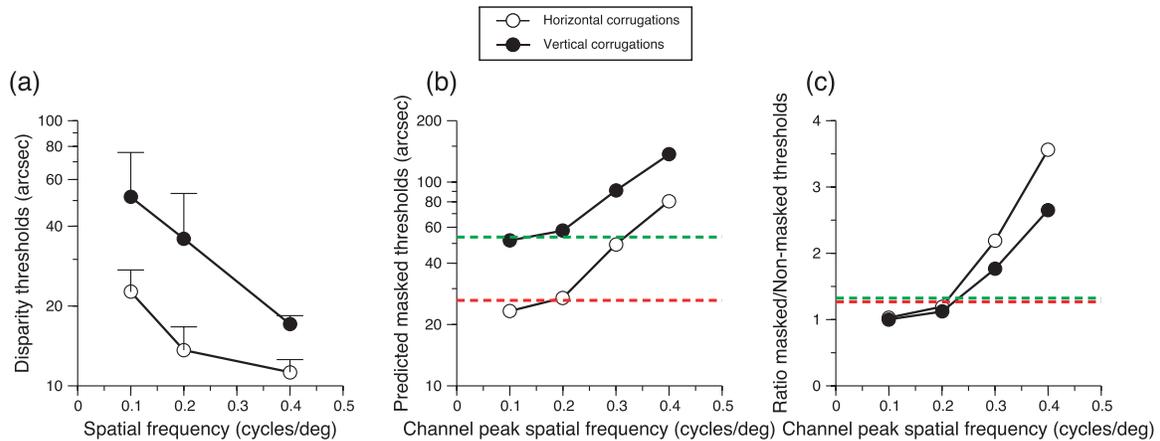


Figure 6. Predictions of the model. White circles, horizontal corrugations; black dots, vertical corrugations. (a) Disparity thresholds (arcsec) as a function of the spatial frequency of the sinusoidal corrugation (mean + SEM, for spatial frequencies 0.1 and 0.2 cycles/° we tested eight subjects, for 0.4 cycles/° we tested four subjects). (b) Predicted masking disparity thresholds (arcsec) as a function of the peak of the channel that detects the sinusoidal corrugation of spatial frequency of 0.1 cycles/° masked by an ideal 1D band-pass noise centered in 0.4 cycles/° and half octave wide. We assumed that the visual channels have the same spatial-frequency bandwidths and those were taken from Figure 4). The red dashed line shows the mean (from eight subjects) of the masked thresholds for horizontal corrugations; and the green dashed line shows the mean (from eight subjects) for vertical corrugations. The noise level (power spectral density) for the band-pass noise was  $25 \times 10^{-3}$  (cycles/°)<sup>-1</sup>. (c) Ratios of the predicted masked thresholds (see Panel b) and the nonmasked disparity thresholds for a sinusoidal corrugation of 0.1 cycles/° (see Panel a) as a function of the peak channel that detects the signal. The red dashed line shows the mean of the empirical ratios of eight subjects for horizontally oriented sinusoidal corrugation of spatial frequency of 0.1 cycles/° without and with masking 1D band-pass noise centered in 0.4 cycles/°. The green dashed line shows the mean ratio for vertical corrugations.

underlying channels, probably accounting for the differences between their results and ours.

The bandwidths we estimate for channels are similar to bandwidths taken from adaptation curves (Schumer & Ganz, 1979). These authors used selective adaptation with horizontally oriented sinusoidal depth corrugations and concluded that stereo vision contains multiple channels each selective to a broad range of horizontally oriented spatial frequencies of disparity modulation (2–3 octaves, full bandwidth at half amplitude). They estimated the bandwidth directly from the threshold elevation plot, assuming that the elevation threshold plot shows the shape of the disparity channel.

In Experiment 2 we measured disparity thresholds for detecting a sinusoidal corrugation of 0.1 cycles/° (horizontally and vertically oriented) under unmasked and masked conditions. For the masked condition we used 1D band-pass noise centered on 0.4 cycles/° and 0.5 octaves wide. Figure 4 shows the disparity thresholds for unmasked and masked conditions. Figure 4 also shows the ratio of the disparity thresholds for both experimental conditions. Using the bandwidths estimated from the average data of Experiment 1, we predicted the disparity thresholds for the masking condition assuming single or multiple channel detection (Figure 6).

Both hypotheses assume that stereo vision possesses a channel centered on 0.4 cycles/°, where human

sensitivity is greatest, whose bandwidth was measured in Experiment 1. The single-channel hypothesis assumes that this is the only channel present. The signal at 0.1 cycles/° would then be detected by this 0.4 channel. Being 2 octaves away from the channel's peak, the signal would be relatively hard to detect. Conversely, noise at 0.4 cycles/° would be where the channel is most sensitive and would thus have a powerful effect. We would therefore expect thresholds for detecting a signal at 0.1 cycles/° to be greatly elevated in the presence of noise at 0.4 cycles/°. We used the power-spectrum model of visual masking to quantify this. For vertical corrugations, the predicted ratio of the thresholds with/without mask was about 2.5 under this single-channel hypothesis (compared to 3.5 for horizontal corrugations). Our experimental data contradict this.

Conversely, the multiple-channel hypothesis assumes that the channel at 0.4 cycles/° is just one of several such channels. In Experiment 1, we estimated the channel bandwidth as no more than 3 octaves, meaning that the channel's modulation transfer function has fallen to half its amplitude at 1.5 octaves from its peak. The noise is centered 2 octaves away from the signal, at 0.4 cycles/° and the lowest frequencies present in the noise are still 1.75 octaves away. Thus, for a channel centered on 0.1 cycles/° noise 2 octaves away has essentially no effect on the detectability of the signal.

Empirically, the ratio for the thresholds with/without masking are close to one, indicating that the human stereo system must contain (at least) two channels for detecting horizontally oriented and vertically oriented depth modulations. Our results do not prove that there is necessarily a channel centered on 0.1 cycles/°. Our quantitative modelling shows that a second channel centered on any frequency up to 0.2 cycles/° would explain our data as well (Figure 6).

These results contradict our previous speculation (Serrano-Pedraza & Read, 2010) that only one channel exists for vertical corrugations. One reason for this speculation was the observation that the disparity sensitivity function is narrower for vertical than for horizontal depth corrugations. We have now measured the bandwidth of disparity channels tuned to horizontal and vertical depth corrugations and found that the bandwidth is consistently slightly narrower for vertically oriented channels. This helps to explain why the disparity sensitivity function is narrower for vertical corrugations, even though it too is made up of at least two channels. Additionally, the fall-off in sensitivity at low frequencies is much steeper for vertical channels than for horizontal channels. We conclude that the stereo anisotropy is due to the poor sensitivity of the lower-frequency vertical channels, not to their absence.

*Keywords:* stereopsis, binocular vision, power-spectrum model of visual masking, orientation stereo anisotropy, linear stereo mechanisms

## Acknowledgments

The findings described here have been reported previously at the Vision Science Society conference (Serrano-Pedraza, Brash, & Read, 2013b). Supported by the Royal Society (University Research Fellowship UF041260 to JCAR) and by grant PSI2011-24491 from Ministerio de Economía y Competitividad (Spain) to ISP.

After this paper was accepted, Prof. Bart Farell pointed out to us that we have not demonstrated noise at 0.4 cycles/° raises thresholds even at 0.4 cycles/°. It is theoretically possible that this noise is simply ineffective as a mask. To address this, we ran a control experiment with author ISP and five further subjects, in which we measured the thresholds for detecting a corrugation at 0.4cpd, either unmasked or masked by noise at 0.4 cycles/° as in Experiment 2. The geometric mean of the ratios for six subjects is 2.9 for Horizontal (range: 1.6–4.6) and 2.1 for Vertical orientations (range: 1.2–3.2). This confirms that the noise stimulus used in Experiment 2 does elevate thresholds for

appropriate stimuli. We are grateful to Prof. Farell for raising this question.

Commercial relationships: none.

Corresponding author: Ignacio Serrano-Pedraza.

Email: iserrano@psi.ucm.es.

Address: Universidad Complutense de Madrid, Faculty of Psychology, Madrid, Spain.

## References

- Anderson, A. J. (2003). Utility of a dynamic termination criterion in the ZEST adaptive threshold method. *Vision Research*, *43*, 165–170.
- Blackwell, K. T. (1998). The effect of white and filtered noise on contrast detection thresholds. *Vision Research*, *38*, 267–280.
- Bradshaw, M. F., Hibbard, P. B., Parton, A. D., Rose, D., & Langley, K. (2006). Surface orientation, modulation frequency and the detection and perception of depth defined by binocular disparity and motion parallax. *Vision Research*, *46*(17), 2636–2644.
- Bradshaw, M. F., & Rogers, B. J. (1999). Sensitivity to horizontal and vertical corrugations defined by binocular disparity. *Vision Research*, *39*(18), 3049–3056.
- Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*(4), 433–436.
- Cagenello, R., & Rogers, B. J. (1993). Anisotropies in the perception of stereoscopic surfaces: The role of orientation disparity. *Vision Research*, *33*, 2189–2201.
- Cobo-Lewis, A. B., & Yeh, Y. (1994). Selectivity of cyclopean masking for the spatial frequency of binocular disparity modulation. *Vision Research*, *34*(5), 607–620.
- Davenport, W. B., & Root, W. L. (1958). An introduction to the theory of random signals and noise. New York, NY: McGraw Hill.
- Emerson, P. L. (1986). Observations on maximum-likelihood and Bayesian methods of forced-choice sequential threshold estimation. *Perception & Psychophysics*, *39*, 151–153.
- Fletcher, H. (1940). Auditory patterns. *Reviews of Modern Physics*, *12*, 1861–1881.
- García-Pérez, M. A. (1998). Forced-choice staircases with fixed steps sizes: Asymptotic and small-sample properties. *Vision Research*, *38*, 1861–1881.
- Green, D. M., & Swets, J. A. (1974). *Signal detection*

- theory and psychophysics* (Reprint with corrections of the original 1966 ed.). Huntington, NY: Robert E. Krieger Publishing Co.
- Guillam, B., & Ryan, C. (1992). Perspective, orientation disparity, and anisotropy in stereoscopic slant perception. *Perception, 21*(4), 427–439.
- Hibbard, P. B., Bradshaw, M. F., Langley, K., & Rogers, B. J. (2002). The stereoscopic anisotropy: Individual differences and underlying mechanisms. *Journal of Experimental Psychology-Human Perception and Performance, 28*(2), 469–476.
- King-Smith, P. E., Grigsby, S. S., Vingrys, A. J., Benes, S. C., & Supowit, A. (1994). Efficient and unbiased modifications of the QUEST threshold method: Theory, simulations, experimental evaluation and practical implementation. *Vision Research, 34*, 885–912.
- Kleiner, M., Brainard, D. H., & Pelli, D. G. (2007). What's new in Psychtoolbox-3? *Perception, 36* (EVP Abstract Supplement).
- Losada, M. A., & Mullen, K. T. (1995). Color and luminance spatial tuning estimated by noise masking in the absence of off-frequency looking. *Journal of the Optical Society of America A, 12*, 250–260.
- Majaj, N. J., Pelli, D. G., Kurshan, P., & Palomares, M. (2002). The role of spatial frequency channels in letter identification. *Vision Research, 42*, 1165–1184.
- Mitchison, G. J., & McKee, S. P. (1990). Mechanisms underlying the anisotropy of stereoscopic tilt perception. *Vision Research, 30*(11), 1781–1791.
- Moore, B. C. J. (1997). *An introduction to the psychology of hearing*. New York: Academic Press.
- Morrone, M. C., & Burr, D. C. (1988). Feature detection in human vision: A phase-dependent energy model. *Proceedings of the Royal Society of London B, 235*, 221–245.
- Mullen, K. T., & Losada, M. A. (1999). The spatial tuning of color and luminance peripheral vision measured with notch filtered noise masking. *Vision Research, 39*, 721–731.
- Nelder, J. A., & Mead, R. (1965). A simplex method for function minimization. *Computer Journal, 7*, 308–313.
- Patterson, R. D. (1976). Auditory filter shapes derived with noise stimuli. *Journal of the Acoustical Society of America, 59*, 640–654.
- Patterson, R. D., & Moore, B. C. J. (1986). Auditory filters and excitation patterns as representations of frequency resolution. In: B. C. J. Moore (Ed.) *Frequency selectivity in hearing* (pp. 123–177). New York: Academic Press.
- Patterson, R. D., & Nimmo-Smith, I. (1980). Off-frequency listening and auditory-filter asymmetry. *Journal of the Acoustical Society of America, 67*, 229–245.
- Pelli, D. G. (1981). *Effects of visual noise*. Unpublished doctoral dissertation, Cambridge University.
- Pelli, D. G. (1997). The VideoToolbox software for visual psychophysics: Transforming numbers into movies. *Spatial Vision, 10*(4), 437–442.
- Pentland, A. (1980). Maximum likelihood estimation: The best PEST. *Perception & Psychophysics, 28*, 377–379.
- Perkins, M. E., & Landy, M. S. (1991). Nonadditivity of masking by narrow-band noises. *Vision Research, 31*(6), 1053–1065.
- Rogers, B. J., & Graham, M. E. (1982). Similarities between motion parallax and stereopsis in human depth perception. *Vision Research, 22*, 261–270.
- Schumer, R., & Ganz, L. (1979). Independent stereoscopic channels for different extents of spatial pooling. *Vision Research, 19*, 1303–1314.
- Serrano-Pedraza, I., Brash, C., & Read, J. C. A. (2013b). Testing the horizontal-vertical stereo anisotropy with the power-spectrum model of visual masking. Vision Sciences Society (Naples, Florida), USA. *Journal of Vision*.
- Serrano-Pedraza, I., & Read, J. C. A. (2009). Stereo vision requires an explicit encoding of vertical disparity. *Journal of Vision, 9*(4):3, 1–13, <http://www.journalofvision.org/content/9/4/3>, doi:10.1167/9.4.3. [PubMed] [Article]
- Serrano-Pedraza, I., & Read, J. C. A. (2010). Multiple channels for horizontal, but only one for vertical corrugations? A new look at the stereo anisotropy. *Journal of Vision, 10*(12):10, 1–11, <http://www.journalofvision.org/content/10/12/10>, doi:10.1167/10.12.10. [PubMed] [Article]
- Serrano-Pedraza, I., & Sierra-Vázquez, V. (2006). The effect of white-noise mask level on sinewave contrast detection thresholds and the critical-band-masking model. *Spanish Journal of Psychology, 9*(2), 249–262.
- Serrano-Pedraza, I., Sierra-Vázquez, V., & Derrington, A.M. (2013a). The power-spectrum model of visual masking: Simulations and empirical data. *Journal of the Optical Society of America A, 30*(6), 1119–1135.
- Solomon, J. A. (2000). Channel selection with non-white-noise mask. *Journal of the Optical Society of America A, 17*, 986–993.
- Solomon, J. A., & Pelli, D. G. (1994). The visual filter

mediating letter identification. *Nature*, 369, 395–397.

- Talgar, C. P., Pelli, D. G., & Carrasco, M. (2004). Covert attention enhances letter identification without affecting channel tuning. *Journal of Vision*, 4(1):3, 22–31, <http://www.journalofvision.org/content/4/1/3>, doi:10.1167/4.1.3. [PubMed] [Article]
- Treutwein, B. (1995). Adaptive psychophysical procedures. *Vision Research*, 35(17), 2503–2522.
- Tyler, C. W. (1975). Stereoscopic tilt and size aftereffects. *Perception*, 4, 187–192.
- Tyler, C. W. (1983). Sensory processing of binocular disparity. In C. Schor & K. J. Ciuffreda (Eds.), *Vergence eye movements: Basic and clinical aspects of binocular* (pp. 199–195). London, UK: Butterworths.
- Tyler, C. W. (1991). Cyclopean vision. In D. Regan (Ed.), *Vision and visual dysfunction*, Vol 9, *Binocular vision* (pp. 38–74). London, UK: Macmillan.
- van der Willigen, R. F., Harmening, W. M., Vossen, S., & Wagner, H. (2010). Disparity sensitivity in man and owl: Psychophysical evidence for equivalent perception of shape-from-stereo. *Journal of Vision*, 10(1):10, 1–11, <http://www.journalofvision.org/content/10/1/10>, doi:10.1167/10.1.10. [PubMed]
- Watson, A. B., & Robson, J. G. (1981). Discrimination at threshold: Labelled detectors in human vision. *Vision Research*, 21, 1115–1122.
- Westrick, Z. M., Henry, C. A., & Landy, M. S. (2013). Inconsistent channel bandwidth estimates suggest winner-take-all nonlinearity in second-order vision. *Vision Research*, 5(81), 58–68.
- Witz, N., & Hess, R. F. (2013). Mechanisms underlying global stereopsis in fovea and periphery. *Vision Research*, 13(87), 10–21.

## Appendix

In this appendix we show the solution of the definite integrals for the assumed modulation transfer function (MTF) for the visual disparity channels. We will also show the particular expressions of the integrals for each noise used in the masking experiments (broadband, notched, and band-pass noise).

### Definite integral of the assumed MTF of the visual disparity channels

The expression of this MTF is presented in Equation 3. We solved the integral within the intervals  $u_{lo}$  (low

spatial frequency) and  $u_{hi}$  (high spatial frequency). This solution is useful in order to solve the integral of the model when we multiply the MTF of the disparity channel by the power spectrum  $\rho$  of the mask (see Equation 3). The MTF is even symmetry ( $H(u, \xi) = H(-u, \xi)$ ) so we need only evaluate the positive half of the integral.

$$\begin{aligned} & \int_{u_{lo}}^{u_{hi}} |H(u; \xi_i)|^2 du \\ &= \int_{u_{lo}}^{u_{hi}} \exp\left[-\frac{\ln^2(u/\xi_i)}{\alpha_i^2}\right] du \\ &= \frac{\sqrt{\pi}\alpha_i\xi_i}{2} \exp\left[\frac{\alpha_i^2}{4}\right] \\ & \quad \times \left\{ \operatorname{erf}\left[\frac{\ln(u_{hi}/\xi_i) - (\alpha_i^2/2)}{\alpha_i}\right] \right. \\ & \quad \left. - \operatorname{erf}\left[\frac{\ln(u_{lo}/\xi_i) - (\alpha_i^2/2)}{\alpha_i}\right] \right\}, \quad (\text{A1}) \end{aligned}$$

where  $u > 0$ ,  $\xi_i > 0$  ( $\xi_i$  corresponds to the peak of the MTF),  $u_{hi} > u_{lo} \geq 0$ ,  $\alpha_i > 0$ , (index of the spatial spread of the MTF of the disparity channel  $i$ ), and  $\operatorname{erf}(x)$  is the error function:  $\operatorname{erf}(x) = (2/\sqrt{\pi})\int_0^x \exp[-t^2]dt$ .

### Model expressions for different masking noises

Here we show the expression of the power-spectrum model for different masking noises.

#### Broadband noise (white noise)

The broadband noise used in Experiment 1 has a constant power density, therefore  $\rho(u) = N_0$ . As explained in the text, we can assume that the signal is detected by the disparity channel tuned to the signal (i.e.,  $\xi_k = u_0$ ,  $|H(u_0; \xi_k)|^2 = 1$ ), therefore, the equation of the model (see Equation 3) used in the fitting is as follows:

$$\begin{aligned} & m_T^2(u_0, u_{lo}, u_{hi}, N_0) \\ &= m_0^2(u_0) + \frac{4N_0}{s} \int_{u_{lo}}^{u_{hi}} |H(u; u_0)|^2 du, \quad (\text{A2}) \end{aligned}$$

where we used  $u_{lo} = 0.04$  cycles/°,  $u_{hi} = 2.5$  cycles/°, and five power spectral density or noise levels,  $N_0 \in \{1.625, 6.5, 25, 100, 400\} \times 10^{-3}$  (cycles/°)<sup>-1</sup>, and  $u_0 = 0.4$  cycles/°. The integral is solved in Equation A1.

#### Notched noise

The notched noise was constructed as the sum of a low-pass noise and a high-pass noise. Again, we have

$\xi_k = u_0$ ,  $|H(u_0; \xi_k)|^2 = 1$ . Therefore, the equation of the model (see Equation 3) used in the fitting is as follows:

$$m_T^2(u_0, u_{lo}, u_{hi}, N_0) = m_0^2(u_0) + \frac{4N_0}{s} \left[ \int_{0.04}^{u_{lo}} |H(u; u_0)|^2 du + \int_{u_{hi}}^{2.5} |H(u; u_0)|^2 du \right], \quad (\text{A3})$$

where  $u_{lo}$  is the cut-off frequency of the low-pass component and  $u_{hi}$  is the cut-off frequency of the high-pass component, and  $u_0 = 0.4$  cycles/°,  $u_{lo} = u_0 2^{-S_{oct}/2}$ , and  $u_{hi} = u_0 2^{S_{oct}/2}$ , where  $S_{oct}$  is the spectral notch size in octaves. The power spectral density of the notched noise was  $N_0 = 25 \times 10^{-3}$  (cycles/°)<sup>-1</sup>.

### Band-pass noise

In Experiment 2 we used band-pass noise of 0.5 octaves around 0.4 cycles/°. The equation of the model (see Equation 2) for this particular mask that we used in the fitting is as follows:

$$m_T^2(u_0, u_{lo}, u_{hi}, N_0) = \frac{m_0^2(\xi_k) + \frac{4N_0}{s} \int_{u_{lo}}^{u_{hi}} |H(u; \xi_k)|^2 du}{|H(u_0; \xi_k)|^2}, \quad (\text{A4})$$

where  $u_0 = 0.1$  cycles/°,  $u_{lo} = 0.336$  cycles/°,  $u_{hi} = 0.475$  cycles/°. The power spectral density of the band-pass noise was  $N_0 = 25 \times 10^{-3}$  (cycles/°)<sup>-1</sup>. For the multiple channel prediction, we used  $\xi_k = 0.1$  cycles/° and for the single channel prediction we used  $\xi_k = 0.4$  cycles/°, as explained in the text.